Economics Letters 156 (2017) 42–47

Contents lists available at ScienceDirect

Economics Letters

journal homepage: www.elsevier.com/locate/ecolet

Conservative vs optimistic rationality in games: A revisitation

Alphonse Fodouop Fotso [a], Roland Pongou [b,*], Bertrand Tchantcho [c]

[a] Department of Mathematics (UY1), Cameroon
[b] Department of Economics, University of Ottawa, Canada
[c] École Normale Supérieure (UY1) and THEMA, Université de Cergy Pontoise, France

HIGHLIGHTS

- Formalize the conservative criterion for complete information games.
- Identify conservative rationality concepts.
- Identify optimistic rationality concepts.
- Find that some optimistic concepts are characterized by irrational optimism.

ARTICLE INFO

Article history:
Received 12 October 2016
Received in revised form 11 February 2017
Accepted 24 February 2017
Available online 20 March 2017

JEL classification:
C7
D8

Keywords:
Sequential games
Conservative criterion
Myopic/farsighted rationality
Regret
Irrational optimism

ABSTRACT

This paper revisits the problem of formalizing a *conservative* rationality concept in games. A rationality concept is said to be conservative if whenever it advises an agent to move from the status quo, that agent cannot be worse off at any of the possible equilibria that may be reached subsequent to this move, relative to the status quo. We formalize this notion for a wide class of games under complete information. Examining some leading concepts of rationality for such games, we find that the only concepts which are conservative are the stable set and the largest consistent set. An implication of our finding is that, along with the Nash equilibrium and the core, Harsanyi's farsighted stable set and most of its descendants in actual fact lack farsightedness. Most of these rationality concepts are premised on a notion of optimism that we find to be "irrational" in that first movers in sequential games do not take into account the fact that subsequent moves are initiated by players who, even if they are optimistic, are utility maximizers.

1. Introduction

This paper revisits the problem of formalizing a *conservative* rationality concept in games. A rationality concept is said to be conservative if none of the possible final outcomes of a move that it recommends to an agent is inferior to the status quo. This property has a natural appeal in the theory of sequential games, where a move by a player or a group of players is likely to trigger further moves by other players, making it essential for players to take into account the long-term consequences of their actions.

The need for a conservative rationality concept partly followed from the observation by several scholars that the classical solution concepts – the stable set (Von Neumann and Morgenstern, 1944), the Nash equilibrium (Nash, 1950), and the core (Gillies, 1959; Aumann, 1961) – are myopic. For instance, Harsanyi (1974) argues that "the von Neumann–Morgenstern definition of stable sets is unsatisfactory because it neglects the destabilizing effect of indirect dominance relations" (p. 1472).[1] He addresses this limitation by introducing the farsighted stable set. In this paper, we revisit this solution concept and some of its most prominent descendants.

As a starting point, we show that while Harsanyi (1974) was correct in noticing that the rationality of the stable set is myopic, his proposed farsighted stable set does not resolve the issue. This point is illustrated by the following example, wherein we show that the rationality of the farsighted stable set – unlike that of the stable set – may recommend actions that agents surely end up regretting.

Example 1. A sequential game involves two players, 1 and 2, and four alternatives $a(0,1)$, $b(1,1)$, $c(2,2)$, and $d(3,0)$, where each alternative is associated with a utility profile. The game is represented by Fig. 1 below.

This graph is interpreted as follows. If a is the status quo, the two players together have the ability to move to b; if they make

* Corresponding author.
E-mail address: rpongou@uottawa.ca (R. Pongou).

[1] Indirect dominance expresses the idea that players can anticipate other players' moves.

http://dx.doi.org/10.1016/j.econlet.2017.02.028
0165-1765/

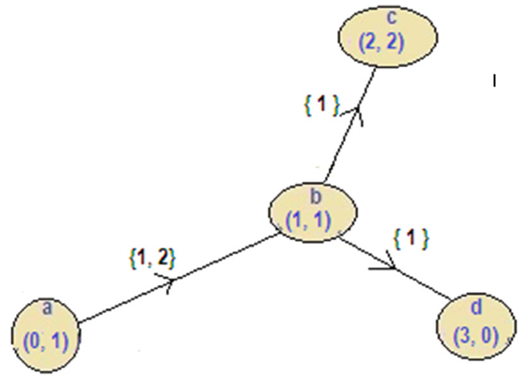

Fig. 1. Analyzing players' behavior under the stable set and the farsighted stable set.

that move, player 1 has the ability to move to c or to d; from c or d, no player or coalition of players has the ability to move.

We now examine the predictions of both the stable set and the farsighted stable set. The stable set of this game is $\{a, c, d\}$, and the farsighted stable set is $\{c, d\}$. This means that, if a is the status quo, following the rationality of the stable set, coalition $\{1, 2\}$ will not replace a by b. This is because no element of the stable set directly dominates a in the sense of being preferred to a by a coalition of players that has the ability to move from a. However, following the rationality of the farsighted stable set, player 2 will join 1 to move from a to b, in the hope that once at b, 1 will move to c (in fact, c indirectly dominates a), which is a better option for 2 than a. But notice that this optimism is "irrational" since at b, a utility-maximizing player 1 will never move to c, given that d is a better option for her than c. It immediately follows that players who follow the rationality of the farsighted stable set are not really "farsighted" since under any reasonable farsighted rule, player 2 should see that at b, player 1 cannot move to c.

The example above shows that in certain situations, players make better decisions by following the rationality of the stable set than by following that of the farsighted stable set. This is because the stable set – while not farsighted in the sense that it does not always prescribe utility-maximizing moves – is conservative, whereas the farsighted stable set is neither farsighted nor conservative.[2]

We also examine important descendants of the farsighted stable set. These include the largest consistent set (Chwe, 1994), the conservative stable standard of behavior, and the optimistic stable standard of behavior (Xue, 1998). More recently, Dutta and Vohra (forthcoming) also introduced the rational and the strong rational expectations farsighted stable sets. Of all these rationality concepts, we find only the largest consistent set to be conservative.[3] An important implication of this finding is that, like the farsighted stable set, the other concepts can be viewed as being premised on a certain notion of optimism in the sense that each of these concepts would sometimes prescribe a deviation from the status quo if at least one predicted outcome of this deviation is superior to the status quo. But for most of these concepts, this optimism is in general naive and consequently irrational in that first movers in general fail to take into account the fact that subsequent moves are made by players who, even if they are optimistic, nevertheless take only actions that maximize their utility.

Table 1
Summary of the findings.

Rationality concept	Is conservative
Nash equilibrium	No
Core	No
Stable sets	Yes
Farsighted stable set	No
Largest consistent set	Yes
Optimistic stable standard of behavior	No
Conservative stable standard of behavior	No
Rational expectations farsighted stable set	No
Strong rational expectations farsighted stable set	No

The rest of this paper is organized as follows. The next section introduces preliminary definitions and formalizes the notion of rationality for a wide class of games. It also formalizes the *conservative* criterion. Section 3 tests the aforementioned solution concepts against this criterion. Section 4 concludes. For clarity, all the proofs are collected in an Appendix.

2. Preliminary definitions

In this section, we introduce preliminary definitions and the key concepts of the paper.

2.1. Games

A game is a situation of interactive decision making which we model as $\Gamma = (N, A, (\to_S)_{S\in P\subseteq 2^N}, (\succsim_i)_{i\in N})$ where N is a finite non-empty set of players; A is a non-empty set of alternatives; P is a collection of admissible coalitions; $(\to_S)_{S\in P\subseteq 2^N}$ is an effectivity function that describes the distribution of ability among the admissible coalitions – where for each admissible coalition $S \in P$, $\to_S$ is a binary relation defined on A –, and $(\succsim_i)_{i\in N}$ is a preference profile. The set 2^N denotes the set of non-empty subsets of N. A game is said to be *finite* if its set of alternatives is finite.

For each admissible coalition S, the effectivity or ability function $\to_S$ determines the set of actions that S can take. For example, for any alternatives a and b, we write $a \to_S b$ to indicate that S has the ability to replace a by b if given the opportunity to do so.

The class of games analyzed here nests several important subclasses, including non-cooperative games, transferable-utility games, network games, and political games. For example, in a classical non-cooperative game, the set P of admissible coalitions is the collection of all the singletons of 2^N. In a bilateral network game, P consists of one-player and two-player sets. In a simple (or political) game (Shapley, 1962), P consists of winning coalitions.[4]

For any individual $i \in N$, the preference relation $\succsim_i$ is assumed to be a complete and transitive binary relation on the set of alternatives A. For any alternatives $a, b \in A$, $a \succsim_i b$ means that i prefers a to b. We denote respectively by $\succ_i$ and $\sim_i$ the asymmetric and the symmetric components of $\succsim_i$.

For any set $S \in 2^N$ of players, we write $a \succsim_S b$ to mean that each individual in S prefers a to b: $a \succsim_i b \ \forall i \in S$. Also, we write $a \succ_S b$ to mean that each individual in S strictly prefers a to b: $a \succ_i b \ \forall i \in S$. We write $a \succ b$ to mean that there exists a set of players $S \in P$ satisfying $b \to_S a$ and $a \succ_S b$.

Let $i \in N$ be a player, $S \in 2^N$ a set of players, $a, b \in A$ two alternatives, and B a subset of A. We define by $\min_i B = \{b \in B : x \succsim_i b \text{ for all } x \in B\}$ the set of the minimal elements of i's preference relation $\succsim_i$ with respect to B. We write $B \succsim_S x$ to mean that for all $b \in \min_i B$, $b \succsim_S x$ and there exists some $b \in B$ such that $b \succ_S x$.

2 The fact that the rationality of the stable set is conservative is surprising because the direct dominance relation that underlies it is myopic. In results presented in the working paper version of the current article (Fodouop et al., 2017a), we show that the Nash equilibrium and the core are based on the same myopic dominance relation, but unlike the stable set, these solution concepts are not conservative. These findings are summarized in Table 1.

3 Herings et al. (2004) also argue that the conservative stable standard of behavior and the optimistic stable standard of behavior may rule out too little; they propose a notion of rationalizability which identifies coalitions likely to form and the alternatives likely to be reached from a given status quo.

4 Simple games are also a subclass of games in constitutional form (Andjiga and Moulen, 1989).